

On Applications of Space–Time Modelling with Open-Source ^{14}C Age Calibration

T. Rowan McLaughlin¹ 

Published online: 28 May 2018
© The Author(s) 2018

Abstract In archaeology, the meta-analysis of scientific dating information plays an ever-increasing role. A common thread among many recent studies contributing to this has been the development of bespoke software for summarizing and synthesizing data, identifying significant patterns therein. These are reviewed in this paper, which also contains open-source scripts for calibrating radiocarbon dates and modelling them in space and time, using the R computer language and GRASS GIS. The case studies that undertake new analysis of archaeological data are (1) the spread of the Neolithic in Europe, (2) economic consequences of the Great Famine and Black Death in the fourteenth century Britain and Ireland and (3) the role of climate change in influencing cultural change in the Late Bronze Age/Early Iron Age Ireland. These case studies exemplify an emerging trend in archaeology, which is quickly re-defining itself as a data-driven discipline at the intersection of science and the humanities.

Keywords Data · Chronology · Radiocarbon · ^{14}C · Calibration · R · Modelling · Simulation · Ireland

Introduction

Archaeology is witnessing a proliferation of studies that avail of large archaeological–chronological datasets, especially radiocarbon data. For example, at the time of writing, recent months have seen papers presenting syntheses of millennia of human history in Central Italy (Palmisano et al. 2017), the Balkans (Porčić et al. 2016),

Electronic supplementary material The online version of this article (<https://doi.org/10.1007/s10816-018-9381-3>) contains supplementary material, which is available to authorized users.

✉ T. Rowan McLaughlin
r.mclaughlin@qub.ac.uk

¹ School of Natural and Built Environment, Queen’s University Belfast, Belfast BT7 1NN, UK

Britain and Ireland (Bevan et al. 2017), western USA (Zahid et al. 2016), the Pacific Northwest (Edinburgh et al. 2017), Eastern Japan (Crema et al. 2016), and Korea (Oh et al. 2017). These studies build on pioneering continental-scale studies also published in recent years (e.g. Hinz et al. 2012; Peros et al. 2010; Shennan et al. 2013; Timpson et al. 2014; Wang et al. 2014; Williams 2013). The current high level of interest in this kind of work doubtless has many causes, including (a) the fact that available archaeological data have now reached a critical mass, allowing the work to proceed (b) hyper-familiarity by archaeologists with computers and the internet, which now pervade virtually every aspect of modern life, (c) next-generation ancient DNA work which has invigorated interest in demography and population history and (d) a broader shift towards “digital humanities”, positioning archaeology at the interface between scientific research and “Big Data” (for an extended discussion, see Kristiansen 2014).

In this brave new world, many archaeologists and their collaborators have developed novel methods and theories dealing with the rather complex and intertwined issues of meaningfully summarizing large datasets, but also accounting for the chronological uncertainty and sampling bias inherent to archaeological research. The ground-breaking nature of much of this effort has led to solutions that are somewhat experimental in nature, often tailored to individual problems, or informed by quite rigid theoretical assumptions. These factors hamper their universal applicability, and, to date, only a small number of platforms for dealing with archaeological-chronological data have been widely adopted (e.g. Bronk Ramsey 2001), although this situation is likely to change (Crema et al. 2017). In this paper, I hope to contribute to this emergent and creative milieu of some universally applicable tools for calibrating radiocarbon dates, examining trends in the data, and mapping out patterns in space and time. Some of the techniques are new, whereas others have been pioneered elsewhere and reviewed here. Also, in this paper and in the digital supplemental appendix, the steps required to perform all these analyses are outlined and entirely coded using the free and open-source platforms R and GRASS GIS, meaning that anyone with a computer can take this work further or adapt it for the needs of their own chronological problem.

Methods for Spatiotemporal Modelling in Archaeology

Radiocarbon Calibration

The calibration process—the calculation of the probability density distributions of radiocarbon dates—is a relatively simple undertaking. To do it, a portion of the calibration dataset or “curve” is selected and used as parameters in the standard equation that describes the probability density of the normal distribution (or alternatively, the t distribution). The resulting set of numbers is usually normalised so that the numerical values correspond to the correct boundary conditions of the probability density function of the “date”. Further details and the R code necessary to calibrate radiocarbon dates are given in the [supplementary materials](#) appended to this paper. This includes a method for calculating the date span of the confidence intervals based on the highest density region of the calibration.

Point Estimates and Sampling from a ^{14}C Calibration Probability Distribution

When working with large ^{14}C datasets or age-depth models, it is sometimes useful to characterise each radiocarbon date by an average value, a “best guess” at a single point in time that describes the samples’ age. A common application is the selection of a subset of dates from a larger database for more rigorous assessment using Bayesian phase models or one of the techniques described in the following texts. Mid-points of confidence intervals or modal dates can be used, or the weighted mean (\bar{x}), which Telford et al. (2004) argue is the best guess of its kind, albeit still a problematic one. Details of how to calculate the weighted mean are contained in the [supplementary materials](#). However, it should be stressed that because the radiocarbon probability density of any calendar year is very low (typically never greater than 0.05), any point estimate of a radiocarbon date is much more likely to be “wrong” than “right”, underlining the need to work with the full probability distribution wherever possible. That said, many studies still succumb to using point estimates or even histograms of uncalibrated dates because there are few methods that can replace the point estimate when the desired result of the analysis is a time-series, such as temporal frequency distribution or plots of proxy indicators versus time.

Bootstrapping is a potentially important application of point estimation that has not yet been widely discussed in the literature (although see Brown 2017) and can address this issue. The technique involves randomly “Monte Carlo” sampling one point from each probability distribution of each ^{14}C date in a set of dates, then using these samples as data in subsequent analyses. Repeating the process many times (*e.g.* 1000), the bootstrapping process develops a set of results that incorporate the error margin caused by the calibration process. Later in this paper, the technique is used to estimate a conservative confidence interval for models of the frequency of ^{14}C dates; it is also a flexible technique and can be used to model ^{14}C -dated processes through time, such as multiple regressions of palaeodietary or geochemical changes associated with a set of dates.

Summed Probability

A summed radiocarbon probability function combines two or more probability density functions for a given date range by adding together their probability densities for each year. As long as each contributing probability density is correctly normalised, the area under the summed probability is equal to the number of items that have been summed. In this way, a time series is constructed that contains information about trends in the frequency of radiocarbon dates, which in turn becomes interpreted as a proxy measurement of past levels of activity. As many authors have pointed out, using the technique uncritically and as a direct proxy is applicable only to broad trends in very large datasets (Chiverrell et al. 2011; Contreras and Meadows 2014; Williams 2012). This is because the inherently statistical nature of radiocarbon measurements, together with the non-Gaussian uncertainty introduced by the calibration process, causes artefacts in the resulting curve that could be misinterpreted as “signal” but are, in fact, “noise”. Defence of the technique has been maintained by many authors, however, as cumulative probability distributions can be seen as a useful way to explore trends in data aggregated from different sites, both in terms of environmental history (*e.g.*

Macklin et al. 2011), cultural change (e.g. McLaughlin et al. 2016), and, when taphonomy can be controlled and the significance of the patterns tested, past demographics (e.g. Crema et al. 2016; Porčić et al. 2016; Shennan et al. 2013).

A number of recent papers attempt to improve summed probability as a modelling technique by confronting criticisms that they are unprobabilistic representations of research interest, not a true pattern borne from fluctuating activity levels in the past. Gaining momentum is what Edinborough et al. (2017) call the “University College London (UCL) method”, where simulation and back-calibration are used to generate a null hypothesis, which can also factor the expected rates of both population growth and taphonomic loss of archaeological materials through time. Through comparison of summed radiocarbon to the simulated model, signals of activity that are significantly weaker or stronger than usual can be identified.

Summed probability distributions can also be used purely as a visualisation tool; a means to explore the temporal relationship of sites and plot the overall distributions of radiocarbon dates on a time series. This allows for the direct comparison of archaeological data to other series and the coarse distribution of sites and their various phases. It is a very useful tool for visualising the latter and probably deserves to be more widely used for exploratory work than it currently is. Plots of the summed probability can suggest patterns of overlap or hiatus and quickly identify chronological patterns that can be investigated further and more rigorously, through Bayesian modelling of apparent cultural phases for example, thereby aiming to make robust estimations of their likely start and end points (Bronk Ramsey 2009).

Kernel Density Estimation

Kernel density estimation (KDE) is an alternative to summed probability modelling. This is a standard statistical technique that aims to smooth out noise in a sample, so that information about the underlying population can be revealed. In the context of radiocarbon data, rather than just capturing and combining information about the uncertainty inherent to the individual samples (*i.e.* measurement error and calibration effects), KDE also attempts to model the uncertainty associated with the sampling procedure itself (e.g. the archaeological excavation). This uncertainty is usually modelled as a Gaussian distribution; other kernels can be used although the Gaussian is arguably the most appropriate way to model the sampling error. The choice of smoothing bandwidth is important—if it is too narrow, then spurious wiggles caused by the calibration curve (or rather, the ^{14}C history of the Earth) will be present; too wide, and the KDE will fail to respond to patterns of interest in the data. The factors determining the choice of kernel bandwidth are, in essence, similar to the measures taken to mitigate the risk of type I and type II errors in research design: both false positives and false negatives should be avoided if possible, but sometimes a study must run the risk of making one type of error to lower the chance of making a more serious error of the other kind. There is no universal solution or consensus; the heuristic of 200 years suggested by e.g. Weninger et al. (2015) as a natural limit to the present resolution of radiocarbon data trends is sufficient for long-term research in prehistoric archaeology, but unlikely to be sufficient for protohistoric work, for example, where changes of less than 100 years are typically of interest. Even in prehistory, the success of Bayesian chronologies in examining individual sites or certain cultural developments has transformed expectations of how

much precision to expect from radiocarbon data. However, because of the spurious and unpredictable nature of calibration noise, it is not appropriate to use KDEs for bandwidths of less than 200 years unless uncertainty from the calibration curve can be factored into the model (*e.g.* Bronk Ramsey 2017, see also discussion in the following texts). Another contention is the choice of point estimate for each radiocarbon date. As mentioned previously, the weighted mean is the best candidate although it remains effected by certain idiosyncrasies of the calibration curve, especially plateaux.

The steps required to compute a KDE time series of radiocarbon data are therefore the following: (a) choose a point from the calibrated probability density of each radiocarbon date, (b) use these as the data parameter in a KDE software function and (c) plot or otherwise analyse the results. In the subsequent texts, I suggest a method for factoring in the calibration uncertainty using a bootstrapping random sampling method. This produces results that are comparable to the more complicated and robust method proposed recently by Bronk Ramsey (2017) and is much faster to run and therefore useful for exploratory work. Readers are encouraged to also look to Bronk Ramsey’s method and draw comparisons between it and the tools presented here.

Suggested Refinements: Confidence Intervals and Permutation

Many time series-based studies of chronological data struggle to define the degree of confidence associated with the patterns revealed by the analysis, or if they do, the confidence intervals are highly conservative because they are calculated using random permutations of summed probability models and are affected by noise from the calibration dataset. KDE time series suppress noise, but for all practical purposes, their confidence intervals cannot be calculated directly, especially for the complex patterns of uncertainty associated with radiocarbon data. One solution, which I introduce here, is to use a Monte Carlo process to estimate so-called “bootstrap” uncertainty. The method I propose is as follows, with the relevant R code appearing in the supplementary materials:

1. A large catalogue of radiocarbon dates (ideally several thousand entries per millennia) relevant to the research question is obtained and edited such that no site phases are over-represented.
2. A proportion of the entries are randomly selected.
3. Each ^{14}C date is calibrated.
4. From each calibrated probability density function, one calendar year is sampled.
5. The Gaussian kernel density of these sampled ages is estimated.
6. Steps 3–5 are repeated many times (*e.g.* 1000).
7. Steps 2–6 are repeated a few times, to check that the input data is sufficiently resolved.

A model of the changing activity levels in the input data emerges as iterations of the process play out. These can be plotted or summarized using the mean and standard deviation of the Gaussian kernel density of each year, which can be computed “on-the-fly” with streaming input data, after each iteration of step 5.

The choice of bandwidth depends partly on the sample size, as the bootstrap confidence interval will be very large if the sample size is small, and also whether long-term or short-term trends are of interest. For most archaeological applications,

relatively small bandwidths (e.g. 30 years) will probably be preferred as these will respond to sudden changes in the archaeological record. However, if the overall aim is to model evolutionary trends over many millennia, then bandwidths spanning several centuries can be used. As a general rule, if using a smaller bandwidth to attempt to resolve short-term patterns, then the uncertainty associated with the KDE model increases. This is an archaeological “uncertainty principle” of sorts; one can have a highly resolved radiocarbon density model, or one with a narrow confidence interval, but not both.

The random sampling in steps 2 and 7 are optional but partly address the oft-levied criticism that some structure in radiocarbon time series could be caused by the incompleteness or bias of the input data. If similar results are obtained from different subsets of the input data, then this constitutes evidence that the sample size (*i.e.* the number of dates) is sufficient. Alternatively, if the results of several different permutations do not converge on a single pattern, then it is likely that the sample size is insufficient. Varying the number of random samples that constitutes each permutation is also a useful way of exploring the minimum number of radiocarbon dates needed to resolve patterns—this number is likely to depend on the magnitude of the change, the rate of change, and the overall whereabouts of the change in time. For small catalogues of dates however, it would be appropriate to omit these steps completely.

Drawing inspiration from the “modified UCL” method (Edinborough et al. 2017) in summed radiocarbon studies, it is possible to set up a null hypothesis model using bootstrapped radiocarbon Gaussian KDE models calculated using the above-mentioned steps and formally test whether the archaeological data significantly deviate from a trajectory that would be expected if all things were equal. To do this, the process is run on both simulated radiocarbon data (the null hypothesis) and the archaeological data (the radiocarbon date catalogue). The kernel densities, which were computed in step 5 for each time coordinate, are used as the “data” in such tests. The R code needed to carry this out, including an easy-to-use function `MCdensity`, can be found in the [supplementary materials](#). A comparison between the output of `MCdensity` and `OxCal 4.3.2` (Bronk Ramsey 2017) appears in the [supplementary information](#).

It is worthwhile comparing the bootstrapped KDE models with summed probability models using the same set of simulated data in Fig. 1. In this exercise, a hypothetical model was defined describing declining activity spanning AD 400 and 650 followed firstly by steady activity levels, then a short increase in activity. This was used to “back-calibrate” a set of simulated 727 radiocarbon “dates” that could be subjected to both forms of analysis (following Timpson et al. 2014). The resulting KDE (Fig. 1a) reflects well the original model, aside from inevitable edge effects that hamper the detection of the short activity spike around AD 1050. The summed probability model (Fig. 1b) by contrast contains significant calibration noise, and when a 30-year moving average is used to smooth these out (*cf.* Williams 2013), the model fails to recognise the significant period of decline spanning AD 400 and 650. Moreover, when permutation of the data is used to derive a confidence interval for the summed probability model, the permutations consistently overestimate the activity levels between AD 600 and 1000, which in this case, is likely to have been caused by a slight plateau in the `IntCal13` calibration curve (Reimer et al. 2013; see also discussion in Bronk Ramsey 2017). This exercise illustrates the benefits of the KDE approach, as the calibration noise is restricted to the bootstrap confidence interval, whereas the shape of the KDE is

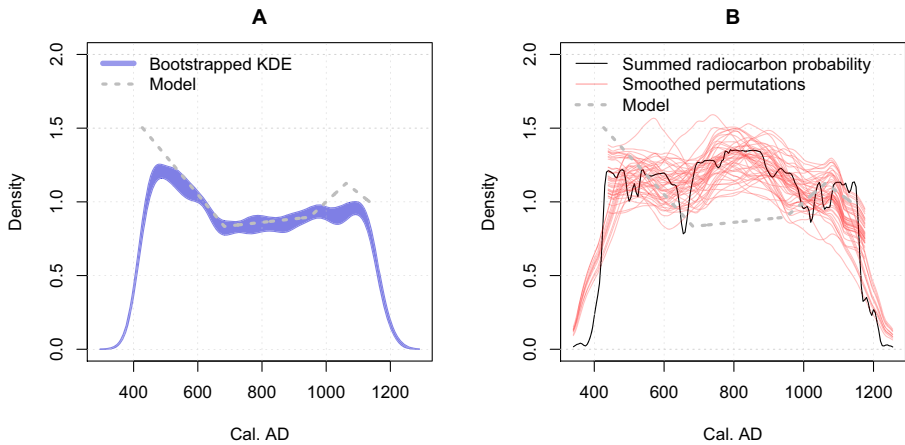


Fig. 1 A comparison of (a) KDE and (b) summed probability models of the same simulated set of 727 radiocarbon dates whose distribution is indicated by the dashed line. The bandwidth/smoothing parameter in both cases is 30 years. The density values of the KDE model have been multiplied by the sample size (727) to render them directly comparable with the summed probability.

dependant on the data, with its kernel aiming to smooth over bias and missing elements. Included in the [supplementary materials](#) within this paper is the R code necessary to simulate sets of radiocarbon dates and analyse them using either a KDE or summed probability analysis for any other point in time where there exists a calibration curve.

Testing the Method: The Black Death

To test whether the bootstrapped KDE method works for a simulation of radiocarbon data informed by real-world parameters, I now turn to a familiar case study, namely the population history of Europe during the Black Death. This was one of the case studies used by Contreras and Meadows (2014) to illustrate the difficulties in identifying such events in radiocarbon data; they concluded that the consequence of the Black Death, a 30% drop in population, is difficult to “see” even when randomly sampling calendar dates from the population curve, and virtually impossible when viewed through the heavy fog of the calibration process. However, as Edinborough et al. (2017) point out, such demographic events can be positively associated with changes in simulated summed radiocarbon probability distributions by deploying a hypothesis-testing approach—if, of course, there really is a direct link between demography and the production rate of materials destined to become ^{14}C samples, which is another debate entirely. To test whether the Gaussian KDE can resolve the Black Death “event” under a similar set of assumptions, the historic demographic data were used to simulate various numbers of calendar dates whose frequency distribution in time was randomly sampled from the demographic data supplied by Contreras and Meadows (2014) for the period AD 1000 to 1700. These were “back-calibrated” and re-calibrated (Timpson et al. 2014) and used to build a “bootstrapped” Gaussian KDE based on the Monte Carlo process described previously (for the R code, see the [supplementary materials](#)). For this exercise, the radiocarbon dates were simulated several times, and subsequent runs of Monte Carlo sampling and kernel density estimation were undertaken on each

set of simulated dates. The process was repeated for a null model, where population growth was even and not interrupted by the Black Death (Fig. 2).

The results from nine simulations, shown in Fig. 3, indicate that with large sample sizes, the Black Death event is clearly visible in the simulations—or at least, the KDE model of activity in the Black Death simulations is lower than the activity in the null model simulations. These results also hold for some cases of relatively small samples, for example when the number of dates in the simulation was reduced to 500 (15 out of 30). When increased to 3000 dates, all simulations that were attempted (30 in number) revealed the correct pattern. Interestingly, many of the simulations also resolve a period of heightened activity in the thirteenth century, where the historic demography is at a higher level than the null model. Therefore, it is possible to conclude that the “bootstrapped” Gaussian KDE Monte Carlo models of radiocarbon data can resolve phenomena on the scale of the Black Death, and other even more subtle patterns, if the input data are sufficiently resolved. I should stress that there are no rules determining how many are required. In this case, a subtle pattern required 2500 dates for reliable results; changes of greater magnitude or trends occurring over a longer time base would require considerably fewer dates in the kernel density model. KDE modelling is therefore potentially very useful in a range of archaeological studies where radiocarbon datasets or other well-powered chronological data are available from the sites in question.

Long-Term Trends and the Significance of Short-Term Fluctuations

Although the above-mentioned results seem promising, one critical issue exists in the selection of the null hypothesis because it is needed to interpret the data. Without the advantage of historical data as an interpretative guide, it is unlikely that the Black Death signal would be noticed, despite it being significantly and correctly different from a simulated model of sustained growth. Further, in prehistoric contexts, where by definition no historical data exists to construct null models, the only things available are theoretical constraints informed by human population growth rates (*e.g.* Downey et al. 2014; Timpson et al. 2014) and expected taphonomic loss (Surovell et al. 2009). A profitable way forward is to formulate many different kinds of models and see which one best fits the archaeological data or which ones are refuted as not realistically matching the evidence. Virtually all archaeological research progresses in this way;

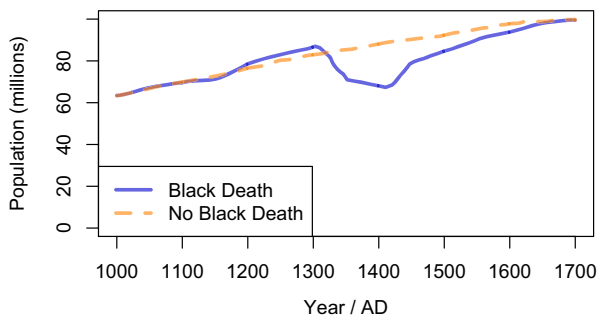


Fig. 2 Population models based on historic data (“Black Death”) and a hypothetical, gradual population increase without any demographic event (“no Black Death”). After Contreras and Meadows (2014)

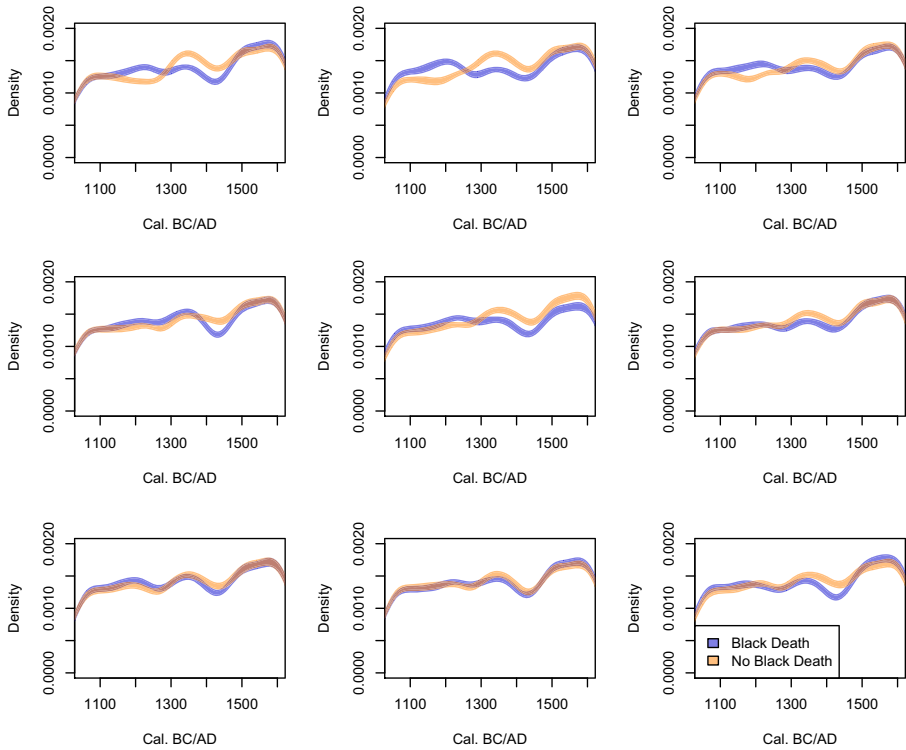


Fig. 3 Nine simulated sets of 2500 ^{14}C dates whose distribution is influenced by the models in Fig. 2 subjected to Monte Carlo KDE analysis. In all but two cases (centre left and bottom left—although the bottom-left case is marginal), the analysis correctly identifies that simulated fourteenth to fifteenth century activity in the “no Black Death” case is higher than the “Black Death” case. In 41 additional runs, one other case was simulated where “Black Death” was not observed. In no cases was the opposite pattern observed

simulating radiocarbon models and testing them with archaeological data is simply an example of how to do this explicitly with respect to large radiocarbon datasets.

An advantage of the KDE method compared to summed radiocarbon is that different models can be calculated using the same input data by changing the bandwidth parameter. For kernels with bandwidths of 100 or more years, what emerges is a picture of long-term change; for 50 years or less, short-term fluctuations become apparent. This presents the possibility of using a KDE with a bandwidth of 150 years for example as a model of long-term trends in the data, and simultaneously using a kernel of 30 years to examine if there were brief episodes where activity was temporarily but significantly lower or higher than the long-term trend. In this way, an interpretation of the data can be achieved without a hypothetical null model. A case study illustrating this, real-world archaeo-chronological evidence of the Black Death, is presented later in the paper.

GIS Analysis

Geographic Information Systems (GIS) provide tools to study the spatial distribution of data and can be used to visualise and analyse chronological information too. Adding

spatial dimensions to the analysis of archaeological chronology can be very revealing. For example, the significance of certain trends is heightened if they can be observed occurring simultaneously and independently in different areas. GIS also offers the possibility to use archaeological dates as models of wider activity in the landscape (Bevan et al. 2013). This is because human activity detected at one site is usually identified as associated with some kind of cultural group, perhaps associated with other broadly contemporary sites in the same region. Through visualising chronological data spatially, interactions between the temporal and spatial properties of the dataset can be observed. Methodologically, achieving such visualisations can be a challenge, usually involving some kind of bespoke software solution or special script, as no GIS had built-in tools for dealing with the peculiarities of archaeological–chronological data, especially radiocarbon data. Studies published to-date tend to be either rasterized approaches, where the radiocarbon dates contribute to a probability surface modelling activity for each grid-square (pixel), or a point-pattern analysis, whereby the size, shape and/or colour of dots on a map are informed by both contextual and chronological factors. The most appropriate solution depends on the nature of the archaeological question and the constraints of the data. In general, rasterized approaches work well for large-scale studies, especially those of remote time periods and geographic areas with low research intensity. Point patterns are most appropriate where the spatial–chronological data are highly resolved and excel where there is a need to also plot the data together with other GIS layers, such as topography or hydrology. Some examples of both approaches are set out at the “[Archaeological Case Studies](#)” section, with further details in the [supplementary information](#).

A key issue in archaeological GIS is that the geographic distribution of archaeological materials reflects first and foremost where fieldwork has taken place. This must be made explicit in the interpretation of the data of course, but to help, there are analytical techniques that attempt to confront this bias. A commonplace technique is some variation of the “relative risk” surface, where first the spatial density of finds or sites is calculated using a 2D Gaussian kernel, then visualised as a ratio of data density to a control surface (e.g. Bevan 2012). For radiocarbon data, Chaput et al. (2015) have devised a relative risk method where the spatial density of all dated sites is used to generate a “sampling intensity map”, a raster control surface whose pixels express the number of nearby sites per unit area. The ratio, calculated for every pixel, of site density in a given time horizon, to the sampling intensity map is then used to map out spatiotemporal changes in population patterns in each time horizon. Comparisons between sequences of these maps are used to identify episodes of migration and change. Indeed, such sequences are particularly effective when viewed as an animation (see the “[Animations](#)” section).

Other studies have attempted to make finer-grained spatiotemporal analyses of smaller regions using Gaussian kernel density weighted with radiocarbon data, for example in Mesolithic Iberia (Grove 2011) or Neolithic Britain (Collard et al. 2010). However, few studies attempt to formally statistically test the patterns produced by GIS analysis of archaeological–chronological data. In a recent publication, Crema et al. (2017) have proposed a spatial permutation test that aims to identify significant hot and cold spots in regions of both space and time. They achieve this by summing the radiocarbon dates for a locality, and for a series of time slices, weighting neighbouring localities *via* a distance decay formula so that the probability of activity in any given

locality is greatly increased if there is also probability of activity at a nearby locality, and weakly increased by activity detected at greater distances. This surface, which they call a locally weighted summed probability distribution, is calculated for many different permutations of the data. Also calculated from the time slices are local growth rates; these are used in the procedure that tests the statistical significance of local change and does so independent of research intensity. Although this intriguing method is not explored by the current study, the R utilities presented as a supplement to this paper are broadly compatible with those of Crema and colleagues, which are also open-sourced, and there are no barriers to readers using both with their data.

Animations

Computer displays offer the opportunity to create and view animated maps that represent evolving changes in the archaeological record. Animated maps remove the need to somehow find a static representation of time, and they appeal to the human visual sense of movement, thereby making apparent subtle features in the data that would otherwise be difficult to perceive. Two main options exist for what to animate; either estimated density surfaces are derived from the distribution of sites at a sequence of time slices, or points are plotted on a map, and can be seen to animate over time. The former solution is a rasterised approach whereas the point-pattern animations are best suited to vectorised GIS.

The methods used to plot animated maps work at any scale. Animations can be a useful way to explore patterns of changing focus at an individual site, or alternatively, they also work at a super-regional scale, modelling large-scale long-term processes, such as colonisation of the Americas (Chaput et al. 2015). The main drawback of animated maps is that they are somewhat technical compared to traditional forms of publication. Some file formats work by default on only some computer systems, which leads end-users to abandon their attempts to view animations. Another issue is that the animation can sweep past the main points of interest all too quickly for the viewer, or too slowly for the parts of the sequence that lack interesting activity. The techniques I showcase in this paper for plotting animated maps solve these issues by using the ubiquitous ISO-standard PDF file format—the animations are simply “pages” that the user flicks through, the animation manifesting as a consequence of changes in the maps from page to page. This way, the user can speed, slow or reverse the animation at will, and zoom in on areas of special interest.

Radiocarbon data present a unique problem for static cartography as there are only limited ways to visualise irregular distributions of probability density, so another advantage of animations is that each time-slice can represent a different probability density. For vector maps, the obvious way to represent this is in varying the size of the point used to plot each date, such that the area of the point is proportional to the calibrated probability density at each time slice (see also McLaughlin et al. 2016). For rasterised Gaussian density maps, the radiocarbon probability density can be used to weight each point used in the estimation of each time slice’s kernel density, *i.e.* increase or decrease its influence in proportion to the radiocarbon probability.

The technical details of how to make animated radiocarbon maps, and those representing other forms of chronological information, are contained in the [supplementary material](#) appended to this paper. The software used is all open-source and freely

available: R (R Core Team 2014), maptools (Bivand and Lewin-Koh 2016) and GRASS GIS (GRASS Development Team 2015).

Archaeological Case Studies

Animating the Spread of Early Neolithic Culture in Europe

An enduring topic of wide interest is the neolithization of Europe, a cultural and economic development of virtually unparalleled significance given subsequent world events and the permanent environmental changes that were a result. This process began when early Neolithic culture moved from the Near East and Anatolia first through Greece around 7000 BC, then *via* the Balkans and peninsular Italy to the interior and western edge of the continent. Much of central Europe was colonised by a distinctive cultural group known as the “LBK” around 5500 BC—an important few centuries that also saw the arrival of the Neolithic in Iberia and intensive activity throughout the Mediterranean region. The Neolithic eventually spread to Britain and Southern Scandinavia shortly after 4000 BC. The remarkable time depth to the process has led to a rich tradition of debate regarding the process of cultural and/or demic diffusion. Thanks to the success of recent, well-funded research projects, a large amount of spatial and radiocarbon data relevant to the phenomenon has been drawn together, much of which made freely available on the internet. A simple but visually effective animated radiocarbon map of the phenomenon can be drawn using data supplied (for example) by Pinhasi et al. (2005), who gathered together radiocarbon data addressing the earliest occurrence of Neolithic cultures in European, Anatolian and Near Eastern settings. The database, published as part of their paper, contains 765 radiocarbon dates associated with site locations in latitude/longitude format. Dates with missing error terms were assigned the mean value of ± 90 BP; then, the database was imported into R and the functions described in the [supplementary data](#) were applied.

The example animation (see Fig. 4 and the [supplementary animation file](#)) illustrates the nature of the process; rarely is the progress across the landscape smooth and progressive, instead activity jumps from place to place and appears in quite distant regions rather suddenly. A particularly notable moment occurs around 5500 cal. BC, which sees sudden expansion across several fronts. The intermittent patterns of movement evident throughout the animation call into question a Neolithic “wave of advance”; something that can be modelled in terms of a wavefront moving a certain number of kilometres per year. Instead, a more realistic model is one where people move sporadically, using multiple points of entry into new regions, and travel across both land- and sea-based routes, and “appearing”, archaeologically speaking, in new areas once a certain pattern of behaviour is established—a process that may take more than a few years (Drake et al. 2016). An alternative explanation, and a clear limitation of the conclusions that may be drawn from this case study (where “only” 765 dates span a 6000-year period in Europe and the Near East), is that the sudden appearance of the Neolithic in new and distant regions could be due to the poorly powered dataset. Discussed subsequently in the “*Fulachtaí fia* (Burnt Mounds) in Ireland” section is the application of spatial kernel density estimation, which attempts to address this problem by interpolating between the known points.

It remains however, in broad terms, that the large-scale pattern caused by the movement of Neolithic culture through the landscape is very evident in the animation. The animation demonstrates how fairly large and complex radiocarbon datasets can be summarized, and the method can prove equally useful for those seeking a way to identify gaps in our knowledge of assess whether a dataset is adequate to address the research question in hand.

Famine and Black Death in Medieval Britain and Ireland

The use of the “Black Death” 30% population reduction in Europe as a case study in radiocarbon simulation (the “[Testing the Method: The Black Death](#)” section) prompts an evaluation of the real-world radiocarbon evidence of this period. I turn to Britain and Ireland specifically, whose dense history in the twelfth to fourteenth centuries portrays a turbulent, dynamic time. From the Norman invasion of 1066 and onwards, centuries of expansion and consolidation of political power followed, and although punctuated by

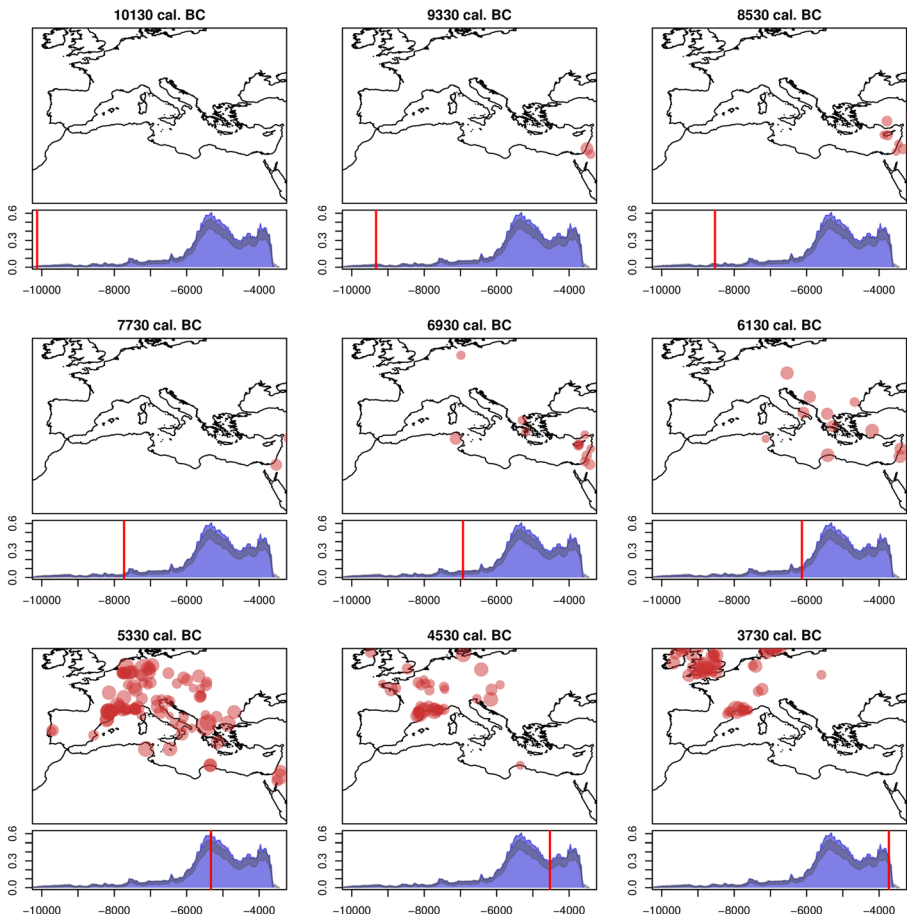


Fig. 4 Example frames from an animated map of the spread of farming in Early Neolithic Europe, plotted using data from Pinhasi et al. (2005) and the R code supplied in the [supplementary materials](#)

many minor wars and rebellions, by the twelfth century, a degree of political and economic stability was established in Britain and Ireland. During this period, there was a move from medieval feudalism towards recognisably modern democracy, with the agreement of *Magna Carta* in 1215, and later in the century, the establishment of parliament under Edward I. However, war between Scotland and England broke out in the 1280s, as did hostilities in Ireland, where Norman control had never become fully established and would quickly wane. In Europe, over-population and inefficient agricultural practices led to demographic crisis and famine early in the fourteenth century. This spread to Britain and Ireland, both directly and through restrictions on the amount of food and other goods that could be traded. Therefore, when the pandemic of plague known as the Black Death struck in 1349, the population of the islands had already been in decline for several generations, and the political and economic systems that could otherwise mitigate its effects had been weakened.

Archaeological data have much to add to knowledge of this period; of particular value are data concerning the response of subsistence-level agriculture to these changing circumstances, as such details are difficult to glean from historic records alone. Although archaeological research traditions for the period have focused on military and ecclesiastic sites, a large number of radiocarbon dates from rural medieval settlements are now available, thanks especially to modern-day infrastructural development and the random sampling of the archaeological landscape carried out as a consequence. Following the results of the simulation exercise presented previously (the “[Testing the Method: The Black Death](#)” section), it can be demonstrated that radiocarbon data have the potential to resolve informative temporal patterns during these centuries. Real-world archaeological data from Britain and Ireland in this period form an ideal case study to explore the effectiveness of the methods discussed earlier in this paper and test whether new information about this critical point in history can be revealed.

And so, a database of 2600 radiocarbon dates for the period AD 1000 to 1700 (included here as [supplementary data](#)) was subjected to a kernel density analysis. The database was compiled through interrogation of existing radiocarbon databases and reviews of publications and grey literature. This review was not exhaustive but compares well with the number of dates compiled independently for the earlier part of this period by another recent study (Bevan et al. 2017), and, therefore, can be considered a representative sample of the data that potentially exists. Reading through the literature, there is an impression that quite a few of the samples in the database are from what were originally thought to be earlier contexts, but upon analysis proved to be medieval in date (quantification of this phenomenon was not attempted). Although perhaps something of a disappointment for the archaeologists who originally submitted these samples, the resulting data is actually particularly valuable for examining demographic and economic trends in the past due to their stochastic nature—results discovered by accident have less bias.

Plots of the Gaussian kernel radiocarbon density computed using the Monte Carlo method in the “[Suggested Refinements: Confidence Intervals and Permutation](#)” section appear in Fig. 4. The dominant trend in these data is caused by the level of interest—or relative lack thereof—in radiocarbon dating the period. Although widely used in Britain and Ireland in prehistory and in early medieval archaeology, radiocarbon dating is less widely used in medieval studies because it lacks the accuracy needed for research questions pertinent to the period. The long-term trend of declining activity

in radiocarbon dates from medieval Britain and Ireland is a serious bias, but one that can readily be modelled using a 150-year kernel bandwidth (Fig. 5, left). By comparing changes in the 30-year KDE to the 150-year KDE, patterns of significantly higher and lower activity levels manifest as departures between the two curves. Although from Britain there is no significant departure from the long-term trend until the late fifteenth century, in Ireland, the data suggest somewhat fluctuating activity levels during the medieval period. These include the historically attested increase of activity in the thirteenth century and a decline in the fourteenth century—a signal perhaps of the European famine and the Black Death reflected in radiocarbon data.

The reason why the event seems present in the Irish data but not in the British data could reflect the large number of small-scale rural settlements excavated in advance of development in Ireland; less work of this kind has been carried out in Britain, especially in the rural landscape (Gardiner and Rippon 2009). This again impresses the point that we can only work with the available data—in an ideal world (perhaps Irish archaeologists are particularly fortunate here)—the radiocarbon density model will approach a true representation of population size but even where bias skews the results, useful summaries of the archaeological record can nonetheless be achieved.

This is exemplified when we turn to the kind of samples that have been dated (Fig. 5, right). During the thirteenth century, in Ireland, there was seemingly a boom in cereals, which declined somewhat during the fourteenth century, but still remained high compared to earlier and later centuries. Interestingly, during the fourteenth century, the number of dates from animal bones actually increased slightly, even though the overall

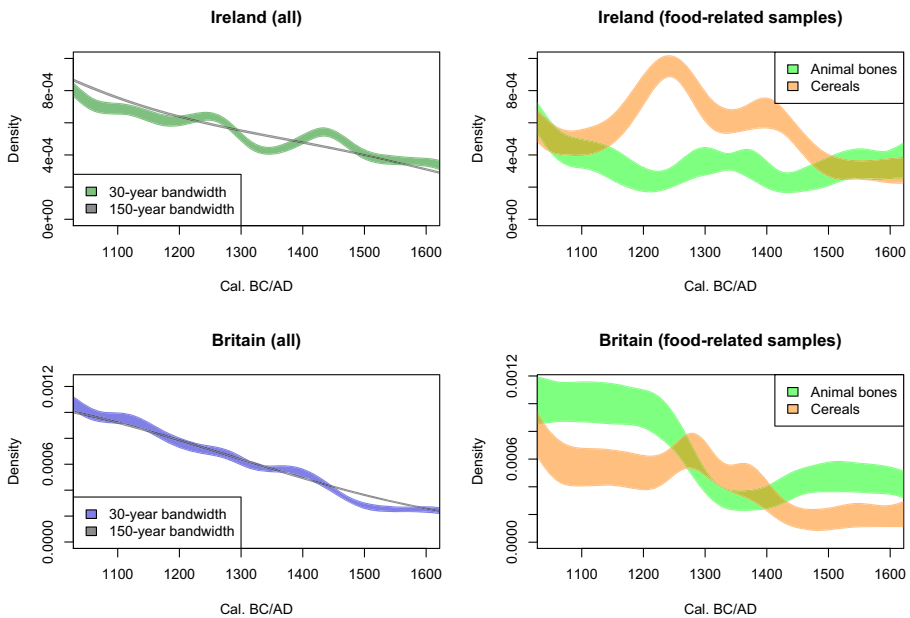


Fig. 5 KDE plots for radiocarbon data from Britain and Ireland in the medieval period showing the differences in long- and short-term trends, and the temporal distribution of ^{14}C dates obtained from charred cereals (Britain, $n = 160$, Ireland $n = 306$) and animal bones (Britain, $n = 259$, Ireland, $n = 208$) (colour figure online)

activity levels (mainly settlements and burials) were depressed. It seems, therefore, in Ireland that the turbulent events of the thirteenth and fourteenth centuries were nonetheless productive times for the island's farmers, even though there was significantly less activity going on in the settlement and burial sites that contribute the bulk of the dates included in the overall density model (Fig. 5, top left). In Britain, a different pattern is evident; the cereals increase whereas the number of dated animal bones is strongly reduced for the fourteenth century. The dearth of fourteenth-century faunal remains in Britain followed a temporary period of increased activity, *i.e.* a kernel density above the long-term trend, which spanned the twelfth and at least the early part of the thirteenth century. In both cases, the radiocarbon evidence reflects the social and economic dynamics of the mid-fourteenth century. Although the number of dates is relatively small, their density through time reveals new information about the past because the underlying changes were of large enough magnitude to be reflected in the data. Had the changes been more subtle, more dates would be needed to resolve any significant pattern in the kernel density models (see the “[Testing the Method: The Black Death](#)” section).

***Fulachtaí fia* (Burnt Mounds) in Ireland**

Fulachtaí fia (Irish Gaelic meaning something comparable to “outdoor cooking places”) are prehistoric burnt mounds (Ó Néill 2003; O’Kelly 1989: 223). Present over much of northern Europe, these sites are ubiquitous on the island of Ireland, and also highly visible archaeologically because they consist of a low mound or spread of burnt stones and charcoal, often covering or beside a “trough”, a subsoil-cut feature. They appear in the landscape in relatively low-lying watery places (Ó Néill 2009). There is little consensus over their interpretation, as in addition to cooking and feasting, other functions such as bathing, dying and brewing have been postulated (*ibid.*). What is apparent, however, is that from their appearance in the Late Neolithic period (Early Neolithic burnt mounds are known too, but these seem culturally and temporally distinctive) until their disappearance in the Early Iron Age, morphologically similar burnt mounds were used on the Island, hinting at a very long-lived pattern of behaviour spanning many other cultural changes and perhaps even at least one population replacement at a genetic level (Cassidy et al. 2016). The frequency that these sites were used can address important questions, namely (a) the periods that they were intensely used, (b) whether there were any hiatuses or fluctuations in-phase with other changes in the archaeological record, and (c) what circumstances brought about the abandonment of this type of site.

To address these questions, a database of 1012 ^{14}C dates from 469 *fulacht* sites was compiled (included as [supplementary data](#)) using publications and archaeological “grey literature”. The majority of these sites (364/469) had only one, two or three associated radiocarbon determinations and only three sites had more than ten—so only very light editing of the data was required to eliminate multiple dates from the same site phase. Much of the data derive from sites discovered during infrastructural development and major building projects. The majority of these dates (798/1012) were from samples of wood charcoal, with others from charred cereals, hazel nutshells, bone and wood. This is a remarkably large ^{14}C dataset for one kind of archaeological feature, although still smaller than (for example) the simulated dataset used to study the Black Death in the

“Testing the Method: The Black Death” section, significant patterns should still be readily quantifiable using the Monte Carlo KDE approach.

The temporal distribution of burnt mounds in Ireland can thus be visualised (Fig. 5) and compared to the dataset for Irish radiocarbon samples provided by Chapple (2015), a subset of which spanning 3500 cal. BC and 500 cal. AD is used here as an indicator of total activity levels from other contexts. The radiocarbon dates from burnt mounds were removed from the comparative dataset, so they are independent of each other, resulting in a sample size of 4394 dates in the comparative dataset. Both datasets show similar peaks in activity at *ca.* 2000, 1500 and 950 cal. BC—the latter peak followed by a swift fall in activity. The KDE plots differ in that the rate of increase in burnt mounds in the late third millennium BC is much greater than the increase in overall total activity, implying an intensifying interest in this type of site during these centuries. This early activity reaches a peak around 2200 cal. BC, which does not correspond to any particular feature of the total activity curve, but does coincide with a proliferation of single burials associated with “food vessel” pottery (Brindley 2007). From this point in time, the plot demonstrates that activity levels at burnt mounds in Ireland were sustained throughout the Bronze Age, and once the tradition became established, the fluctuating levels of activity through time broadly correspond to overall activity levels in all archaeological contexts considered together.

The period 900–700 cal. BC saw a rapid decline in all types of site in Ireland (Armit et al. 2014), including burnt mounds. In the Early Iron Age, between 500 and 400 cal. BC, activity once again increased, but the burnt mound tradition faded entirely at this point in time. A simple explanation for this is that the introduction of new technology and social structure around 500 cal. BC made redundant the burnt mound tradition, which was already much reduced from its apogee centuries before. This is seen in Fig. 6, where the burnt mound data and the overall activity diverge just after 500 BC.

Of particular interest however are the events of 950 to 700 cal. BC. This was a time of rapid environmental and social change in Ireland, although the causality is debated (Armit et al. 2014). A widespread increase in bog surface wetness happened around 750 BC, indicating increased rainfall or lower temperatures, or both. At the most precisely dated site where this is evident, the modelled age range is 786 to 703 cal. BC for the onset of this event (Swindles et al. 2007). The archaeological ^{14}C data certainly suggest rapid decline at this time, but this continues a pre-existing trend. In the KDE plot, there is however a slight interruption of the downward trend between 900 and 800 cal. BC, not dissimilar to the fluctuating activity levels observed since 2000 BC. Using the model building function in the R utilities supplied as supplementary data to simulate, reverse calibrate and re-calibrate various plausible radiocarbon density models (Fig. 7), various explanations for the radiocarbon data can be explored. The best-fit model in Fig. 7 is “Model 1” where activity peaks at 1000 BC declines slightly until 900 BC then remains constant until 750 BC, whereupon it suddenly falls to a low level (simulating a catastrophic reduction in population) then slowly disappears altogether. It should be stressed that the fit does not “prove” the model, as multiple solutions are possible, but it does indicate that the model is consistent with the data, and in this case study, it seems that any conclusions drawn from radiocarbon data that dispute the role of rapid climate change as a potential driver of human population levels (*e.g.* Armit et al. 2014) are premature.

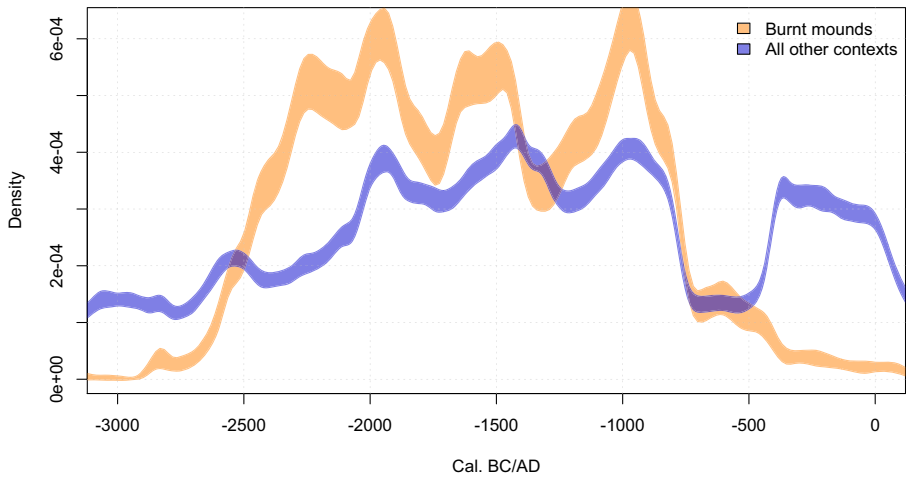


Fig. 6 Monte Carlo ^{14}C KDEs for Irish Burnt mounds ($N=1012$) and a subset of Chapple’s (2018) dataset spanning 3500 BC and AD 500 ($N=4394$)

An analysis of potential spatial aspects to the temporal distribution of burnt mounds in Ireland can be performed using an animated heat map (Fig. 7 and [supplementary materials](#), see also [supplementary materials](#) for how this was constructed). This is quite revealing, illustrating that for the first half of the Bronze Age, the *fulacht* phenomenon was widespread but especially intense in the north of the island, whereas from 1500 to 1400 BC the main focus of activity was in a band across the southern third of Ireland. During lulls in activity, e.g. around 1300 BC, only a small number of sites were in use and even during “booms”, some areas had little or no activity in burnt mounds. Following a decline from 800 to 650 BC, burnt

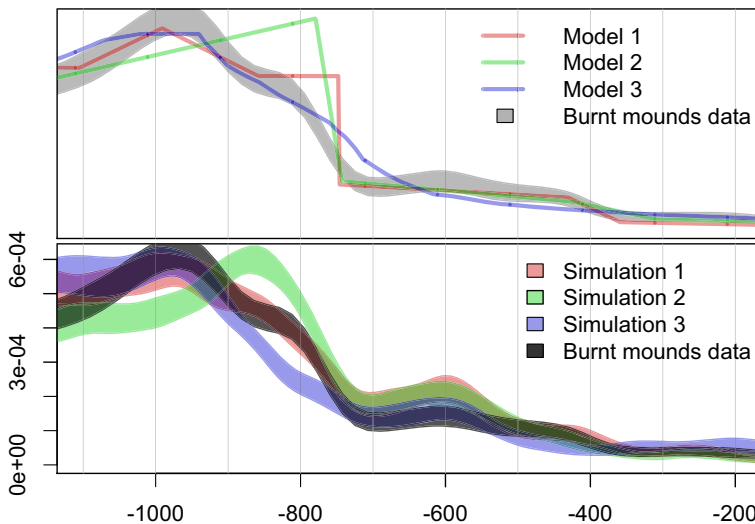


Fig. 7 Simulated ^{14}C Monte Carlo KDEs plotted with real archaeological data from Ireland. Colour figure online



Fig. 8 Gaussian Kernel density maps of *Fulachtai fia* in Ireland at various points in time. Colour figure online

ground activity is taken up in places in the landscape with only limited activity before. This leads to these sites, and any point in time where there are small numbers of dates being quite prominent in the animations. This is because in these maps we are looking at the density relative to the overall pattern throughout time, not the absolute density at any one point in time. This last point illustrates that extensive and intensive patterns are fundamentally different things in archaeology, and most analytical methods can reveal one but not the other; the “relative risk” site density analysis presented here cannot reveal overall intensity, only local intensities, so Fig. 8 should be used with an indicator of the complete picture (*i.e.* Fig. 6).

Summary

A distinct advantage of the Kernel Density approaches in time and space is that the kernel bandwidth is an explicit way to account for bias, missing data, and other issues

with the sampling strategy. It must always be assumed that these factors play a major role in the patterning of archaeological data. Summed radiocarbon analysis does not feature this ability unless subjected to smoothing, which blends together the data and calibration noise in a way that KDE analysis does not. Also, the KDE graphs are straightforward to interpret for archaeological purposes. This is because the choice of kernel bandwidth in KDE analysis reflects the level over which the KDE can reveal trends in terms of how they were experienced in the past; a 30-year bandwidth for example models change at an inter-generational level whereas 150 years a likely span for tradition and cultural memory.

Like any attempt at summary or synthesis, informed critical assessment must be made of the patterns revealed by the results of data analysed using procedures presented here. Regards the animated Gaussian density maps discussed previously, radiocarbon probability was used to influence how 2-D site density was estimated. Although intended as a way to confront the stochastic nature of the data, this method also embeds the maps with a rather complex and debatable notion: activity detected at one coordinate in space–time holds implications for activity levels at other nearby places, immediately before and after the event that was directly dated. In itself, that seems reasonable, as human life is bound to its causes and consequences. However, when using radiocarbon probability density as a model of this influence, the shape of the calibration curve causes some unwanted input to a model that really ought to be completely independent from it. Therefore, the transformation equation or coefficient that maps radiocarbon probability onto the geographical weighting must be carefully designed to avoid undue influence from peaks and plateau in the calibration curve. A future refinement could be to combine the 2-D Gaussian density maps with the “bootstrapped” KDE analysis, which should yield smoother and more conservative distribution maps.

In this paper, I have provided methods and tools that should enable others to summarize radiocarbon data; utilities that aim to extract from data patterns that reflect what we can know about the past, whilst also accounting for what we do not yet know. As neither a mathematician nor a programmer, I cannot hope to have arrived at optimal solutions. Furthermore, long-standing issues, such as research bias and the role of negative evidence are still problematic, and will remain so, as archaeology can only work with what has been unearthed from the ground. It is however pertinent to reflect on how these utilities, and others like them, will be used by archaeologists. As work on the analysis and meta-analysis of archaeological chronology continues apace, two main directions have emerged. The first is centred on Bayesian methods and seeks to answer focussed questions using data from sites that lend themselves to this kind of work or to build age models of events that otherwise could not be dated directly (*e.g.* Whittle et al. 2011). The second kind of study is concerned with large-scale patterns and the meta-analysis of archaeological data from many sites to reveal new and previously unrecognised trends in the human past. The dominant research question demanded during this work has been “what do these data say about human population levels” and this is an extremely important question, potentially contributing to a wider understanding of interactions between human and cultural evolution; albeit one fraught with difficulty (Torfing 2015).

Of enormous value too is a third direction, intermediate in scale, which is in danger of being marginalized with the current focus on other kinds of work. This is the

synthesis: the drawing together of numerous lines of archaeological and other evidence in one cohesive narrative, describing how events unfolded in the past and explaining how the parts of the world came to be. Chronology cannot do this on its own, of course, but it is unrivalled as a coordinating theme, a “golden thread” in an otherwise labyrinthine agglomeration of competing data. To use chronology efficiently, we need flexible tools that can be adapted to the demands of archaeology, the nature of which varies by region and period. The method I suggest here (the “[Suggested Refinements: Confidence Intervals and Permutation](#)” section) for calculating the Gaussian KDE of sets of radiocarbon data whilst remaining sensitive to uncertainty caused by the calibration process offers one such flexible solution. It also is ideal for model building and evaluation; any number of hypothetical time series can be converted into a simulated list of radiocarbon dates and subjected to analysis in order to see what real data might look like. Although this proves nothing in a mathematical sense, as Bevan et al. (2013) put it “a smaller set of equally plausible models is still better than a situation in which anything goes”. Modelling, gathering data and testing the data against the model have been long-standing rudimentary aspects of archaeological research; the opportunity now exists to explore the past through highly resolved data in a way that before seemed a distant dream.

Acknowledgements Maps were plotted with data supplied by the Natural Earth (www.naturalearthdata.com). The radiocarbon data were originally collected by a great many different archaeologists; thanks to the Council for British Archaeology, Historic Environment Scotland, Transport Infrastructure Ireland and Mr. R. M. Chapple for their efforts to bring them together, thus forming the basis of the databases analysed here. Warm thanks to Christina O’ Regan for helpful comments on the text and to Emma Hannah for comments, feedback and testing some of the R utilities presented here. However, errors in the text and shortcomings in the associated software are my responsibility and mine alone.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

- Armit, I., Swindles, G. T., Becker, K., Plunkett, G., & Blaauw, M. (2014). Rapid climate change did not cause population collapse at the end of the European Bronze Age. *Proceedings of the National Academy of Sciences*, *111*(48), 17045–17049. <https://doi.org/10.1073/pnas.1408028111>.
- Bevan, A. (2012). Spatial methods for analysing large-scale artefact inventories. *Antiquity*, *86*(332), 492–506. <https://doi.org/10.1017/S0003598X0006289X>.
- Bevan, A., Crema, E., Li, X., & Palmisano, A. (2013). Intensities, interactions and uncertainties: some new approaches to archaeological distributions. *Computational Approaches to Archaeological Spaces*, 27–52.
- Bevan, A., Colledge, S., Fuller, D., Fyfe, R., Shennan, S., & Stevens, C. (2017). Holocene fluctuations in human population demonstrate repeated links to food production and climate. *Proceedings of the National Academy of Sciences*, *114*(49), E10524–E10531. <https://doi.org/10.1073/pnas.1709190114>.
- Bivand, R., & Lewin-Koh, N. (2016). Maptools: tools for reading and handling spatial objects. <https://CRAN.R-project.org/package=maptools>.
- Brindley, A. (2007). The dating of food vessels and urns in Ireland: Bronze Age studies 7.
- Bronk Ramsey, C. (2001). Development of the radiocarbon calibration program OxCal. *Radiocarbon*, *43*(2A), 355–363.
- Bronk Ramsey, C. (2009). Bayesian analysis of radiocarbon dates. *Radiocarbon*, *51*(1), 337–360.

- Bronk Ramsey, C. (2017). Methods for summarizing radiocarbon datasets. *Radiocarbon*, 59(06), 1–25. <https://doi.org/10.1017/RDC.2017.108>.
- Brown, W. A. 2017. They don't look different, but they're not the same: formal resemblance and interpretive disparity in the construction of temporal frequency distributions. Unpublished manuscript available from arxiv.org/pdf/1708.00535
- Cassidy, L. M., Martiniano, R., Murphy, E. M., Teasdale, M. D., Mallory, J., Hartwell, B., & Bradley, D. G. (2016). Neolithic and Bronze Age migration to Ireland and establishment of the insular Atlantic genome. *Proceedings of the National Academy of Sciences*, 113(2), 368–373.
- Chapple, R. M. 2015. Catalogue of Radiocarbon Determinations & Dendrochronology Dates. Belfast: Oculus Obscura Press.
- Chaput, M. A., Kriesche, B., Betts, M., Martindale, A., Kulik, R., Schmidt, V., & Gajewski, K. (2015). Spatiotemporal distribution of Holocene populations in North America. *Proceedings of the National Academy of Sciences*, 112(39), 12127–12132. <https://doi.org/10.1073/pnas.1505657112>.
- Chiverrell, R. C., Thorndycraft, V. R., & Hoffmann, T. O. (2011). Cumulative probability functions and their role in evaluating the chronology of geomorphological events during the Holocene. *Journal of Quaternary Science*, 26(1), 76–85. <https://doi.org/10.1002/jqs.1428>.
- Collard, M., Edinborough, K., Shennan, S., & Thomas, M. G. (2010). Radiocarbon evidence indicates that migrants introduced farming to Britain. *Journal of Archaeological Science*, 37(4), 866–870.
- Contreras, D. A., & Meadows, J. (2014). Summed radiocarbon calibrations as a population proxy: a critical evaluation using a realistic simulation approach. [article]. *Journal of Archaeological Science*, 52, 591–608. <https://doi.org/10.1016/j.jas.2014.05.030>.
- Crema, E. R., Habu, J., Kobayashi, K., & Madella, M. (2016). Summed probability distribution of ^{14}C dates suggests regional divergences in the population dynamics of the Jomon period in eastern Japan. *PLoS One*, 11(4), e0154809. <https://doi.org/10.1371/journal.pone.0154809>.
- Crema, E. R., Bevan, A., & Shennan, S. (2017). Spatio-temporal approaches to archaeological radiocarbon dates. *Journal of Archaeological Science*, 87, 1–9. <https://doi.org/10.1016/j.jas.2017.09.007>.
- Downey, S. S., Bogaard, E., Kerig, T., Edinborough, K., & Shennan, S. (2014). The Neolithic demographic transition in Europe: correlation with juvenility index supports interpretation of the summed calibrated radiocarbon date probability distribution (SCDPD) as a valid demographic proxy. *PLoS One*, 9(8), e105730. <https://doi.org/10.1371/journal.pone.0105730>.
- Drake, B. L., Blanco-González, A., & Lillios, K. T. (2016). Regional demographic dynamics in the Neolithic transition in Iberia: results from summed calibrated date analysis. *Journal of Archaeological Method and Theory*, 1–17.
- Edinborough, K., Porčić, M., Martindale, A., Brown, T. J., Supramant, K., & Ames, K. M. (2017). Radiocarbon test of demographic events in written and oral history. *Proceedings of the National Academy of Sciences*, 114(47), 12436–12441. <https://doi.org/10.1073/pnas.1713012114>.
- Gardiner, M., and Rippon, S. 2009. Looking to the future of medieval archaeology. In Gilchrist, R. A. & Reynolds, A. (Eds) Reflections; 50 years of medieval archaeology, 1957–2007. Society for Medieval Archaeology, pp. 65–75.
- GRASS Development Team (2015). Geographic resources analysis support system (GRASS) software, version 7.0. Open Source Geospatial Foundation. <http://grass.osgeo.org>.
- Grove, M. (2011). A spatio-temporal kernel method for mapping changes in prehistoric land-use patterns. *Archaeometry*, 53(5), 1012–1030. <https://doi.org/10.1111/j.1475-4754.2010.00578.x>.
- Hinz, M., Feeser, I., Sjögren, K.-G., & Müller, J. (2012). Demography and the intensity of cultural activities: an evaluation of Funnel Breaker Societies (4200–2800 cal BC). *Journal of Archaeological Science*, 39(10), 3331–3340. <https://doi.org/10.1016/j.jas.2012.05.028>.
- Kristiansen, K. (2014). Towards a new paradigm? The third science revolution and its possible consequences in archaeology. *Current Swedish Archaeology*, 22, 11–34.
- Macklin, M. G., Jones, A. F., & Lewin, J. (2011). Comment: cumulative probability functions and their role in evaluating the chronology of geomorphological events during the Holocene. In R. C. Chiverrell, V. R. Thorndycraft, & T. O. Hoffmann (Eds.), *Journal of Quaternary Science* 26: 76–85. *Journal of Quaternary Science*, 26(2), 238–240. <https://doi.org/10.1002/jqs.1478>.
- McLaughlin, T. R., Whitehouse, N. J., Schulting, R. J., McClatchie, M., Barratt, P., & Bogaard, A. (2016). The changing face of Neolithic and Bronze Age Ireland: a big data approach to the settlement and burial records. *Journal of World Prehistory*, 29(2), 117–153. <https://doi.org/10.1007/s10963-016-9093-0>.
- Ó Néill, J. (2003). Lapidibus in igne calefactis coquebatur: the historical burnt mound 'tradition'. *The Journal of Irish Archaeology*, 12(13), 79–85.
- Ó Néill, J. (2009). *Burnt mounds in northern and western Europe: a study of prehistoric technology and society*. Saarbrücken: VDM Verlag Dr. Müller.

- O'Kelly, M. J. (1989). *Early Ireland: an introduction to Irish prehistory*. Cambridge: Cambridge University Press.
- Oh, Y., Conte, M., Kang, S., Kim, J., & Hwang, J. (2017). Population fluctuation and the adoption of food production in prehistoric Korea: using radiocarbon dates as a proxy for population change. *Radiocarbon*, 59(06), 1–10. <https://doi.org/10.1017/RDC.2017.122>.
- Palmisano, A., Bevan, A., & Shennan, S. (2017). Comparing archaeological proxies for long-term population patterns: an example from central Italy. *Journal of Archaeological Science*, 87(Supplement C), 59–72. <https://doi.org/10.1016/j.jas.2017.10.001>.
- Peros, M. C., Munoz, S. E., Gajewski, K., & Viau, A. E. (2010). Prehistoric demography of North America inferred from radiocarbon data. *Journal of Archaeological Science*, 37(3), 656–664. <https://doi.org/10.1016/j.jas.2009.10.029>.
- Pinhasi, R., Fort, J., & Ammerman, A. J. (2005). Tracing the origin and spread of agriculture in Europe. *PLoS Biology*, 3(12), e410. <https://doi.org/10.1371/journal.pbio.0030410>.
- Porčić, M., Blagojević, T., & Stefanović, S. (2016). Demography of the early Neolithic population in central Balkans: population dynamics reconstruction using summed radiocarbon probability distributions. *PLoS One*, 11(8), e0160832.
- R Core Team. (2014). *R: a language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing.
- Reimer, P. J., Bard, E., Bayliss, A., Beck, J. W., Blackwell, P. G., Bronk Ramsey, C., et al. (2013). IntCal13 and Marine13 radiocarbon age calibration curves 0–50,000 years cal BP. *Radiocarbon*, 55(4), 1869–1887. https://doi.org/10.2458/azu_js_rc.55.16947.
- Shennan, S., Downey, S. S., Timpson, A., Edinborough, K., Colledge, S., Kerig, T., et al. (2013). Regional population collapse followed initial agriculture booms in mid-Holocene Europe. *Nature Communications*, 4.
- Surovell, T. A., Byrd Finley, J., Smith, G. M., Brantingham, P. J., & Kelly, R. (2009). Correcting temporal frequency distributions for taphonomic bias. *Journal of Archaeological Science*, 36(8), 1715–1724. <https://doi.org/10.1016/j.jas.2009.03.029>.
- Swindles, G. T., Plunkett, G., & Roe, H. M. (2007). A delayed climatic response to solar forcing at 2800 cal. BP: multiproxy evidence from three Irish peatlands. *The Holocene*, 17(2), 177–182. <https://doi.org/10.1177/0959683607075830>.
- Telford, R. J., Heegaard, E., & Birks, H. J. (2004). The intercept is a poor estimate of a calibrated radiocarbon age. *The Holocene*, 14(2), 296–298.
- Timpson, A., Colledge, S., Crema, E., Edinborough, K., Kerig, T., Manning, K., Thomas, M. G., & Shennan, S. (2014). Reconstructing regional population fluctuations in the European Neolithic using radiocarbon dates: a new case-study using an improved method. *Journal of Archaeological Science*, 52, 549–557.
- Torfiing, T. (2015). Neolithic population and summed probability distribution of 14C-dates. *Journal of Archaeological Science*, 63(Supplement C), 193–198. <https://doi.org/10.1016/j.jas.2015.06.004>.
- Wang, C., Lu, H., Zhang, J., Gu, Z., & He, K. (2014). Prehistoric demographic fluctuations in China inferred from radiocarbon data and their linkage with climate change over the past 50,000 years. *Quaternary Science Reviews*, 98(Supplement C), 45–59. <https://doi.org/10.1016/j.quascirev.2014.05.015>.
- Weninger, B., Clare, L., Jöris, O., Jung, R., & Edinborough, K. (2015). Quantum theory of radiocarbon calibration. *World Archaeology*, 47(4), 543–566. <https://doi.org/10.1080/00438243.2015.1064022>.
- Whittle, A., Healy, F., & Bayliss, A. (2011). *Gathering time: dating the early Neolithic enclosures of southern Britain and Ireland*. Oxford: Oxbow.
- Williams, A. N. (2012). The use of summed radiocarbon probability distributions in archaeology; a review of methods. *Journal of Archaeological Science*, 39(3), 578–589. <https://doi.org/10.1016/j.jas.2011.07.014>.
- Williams, A. N. (2013). A new population curve for prehistoric Australia. *Proceedings of the Royal Society B: Biological Sciences*, 280(1761). <https://doi.org/10.1098/rspb.2013.0486>.
- Zahid, H. J., Robinson, E., & Kelly, R. L. (2016). Agriculture, population growth, and statistical analysis of the radiocarbon record. *Proceedings of the National Academy of Sciences*, 113(4), 931–935. <https://doi.org/10.1073/pnas.1517650112>.